**Papers mentioned by the referees:**

**Ref 1:**

1. **Titman et al (2015)**
2. **Colella et al (2023)**
3. **Reade et al (2022)**
4. **Koszegi & Szeidl (2013) (Behavioral - Theory)**
5. **Shivakumar (2018) (Cricket)**
6. **Chowdhury (2023) (Cricket)**
7. **Ferraresi & Gucciardi (2022)**

**Ref 2:**

1. **Holder et al (2022)**
2. **Oliviera et al (2023)**
3. **Dufner et al (2023)**
4. **Singleton et al (2023)**

**1. Titman et al (2015)**

They talk about how poor referee decisions can ruin the match for a particular team. Example a wrong penalty or red card. The goal of the paper is to model the interdependencies in bookings (red and yellows) in the final outcomes for home and away teams. Actually more than that: what’s the effect of the booking on the course and outcome of a game. They talk about influential events Only use data from 2009-2011 from the EPL and Championship. They conclude reasonable things such as the award of a red card led to increases of goal rates for the non-penalized teams.

I think a good way to cite this paper would be just to mention how important it’s for the ref to have a second opinion and being able to revisit highly influential decisions such as a red car or a goal, things that are scrutinized with the VAR. Cuz according to their model the effect of red cards on scoring appear to differ for the Home and Away teams: Effect of a home red card on home teams scoring: Reduction of probability in 18%. If away team have at least one red card: home team scoring rate increases by 78% substantial. This paragraph might be key:

*The effect of red cards on scoring rates appear to differ for the home (# Home Red) and away(# Away Red) teams. The effect of a home red card on the home team’s scoring rate is relativelymodest (a reduction by 18%). In contrast, if the away team have at least one red card, the hometeam’s scoring rate increases substantially (by 78%). Similarly, if the away team incur a red cardtheir scoring rate drops by 56%. If the home team have a red card this increases the away team’sscoring rate by 62%. Yellow cards appear to have a negligible direct effect on scoring rates foreither team.*

**2. Moral Support and Performance: Colella, Dalton, Giusti (2023)**

They use data from the Argentinean league, when they banned away crowds. They perform a before-after analysis, although they do a counterfactual experiment: The Copa Argentina was played with away crowds.

Still:

*The identification assumption relies on the non existence of other forces that could affect the result of the games and appear contemporaneously with the ban or in the period just after. In other words, we assume that the expected result of every game played before the day in which the law took effect and after that day would be the same if the ban would have never been implemented.*

I think we can cite this paper in our first paragraph, right after describing the 3 main channels:

*One subchannel through which crowds may impact soccer outcomes is by influencing referee decisions; this is known as “referee bias.”*

Colella et al find no evidence of this subchannel, stating that banning away teams fans, doesn’t seem to impact referee’s decisions, based on measures of red and yellow cards and penalties awarded. The mechanism they explain is through moral support rather than pressure to the ref. They test whether banning away crowds increased the hostility (bias?) of referees towards visiting teams and found no significant effect. This conclusion should be taken with some caution, because it seems logical that away crowds would barely influence referees normally, given how especially in south America, local teams only provide a small percentage of the stadium seats to the away crowd.

**3. Reade et al (2022)**

Exploit three changes in either crowded or empty stadiums in Egypt to capture the effect of crowds on behaviour and decision making. The contribution is what they called the underlying dynamics of the effects of crowds on the behaviour of the referees and judges: alternating between periods with and withouth crowds and potentially removing or inducing social pressure!

Interesting that the egyptian context is particularly violent. They mention that changes in the composition of the crowds over the studied period might have changed. Ultra movements fans had been largely dissolved --> reduction in social pressure on the performances of the referee and other officials. Related to that, the paper by Goumas (2014) (Home advantage and referee bias in European football) finds: *After controlling for within-match measures of attacking dominance referees in the Champions League and Europa League issued 25% (p<0.001) and 10% (p=0.002) more yellow cards, respectively, to away teams than to home teams. The higher level of home team bias in the Champions League appeared to be mainly due to higher* ***crowd densities.***

**4. Koszegi & Szeidl (2013) (Behavioral - Theory)**

Richard: I just read the introduction and conclusion to this paper; it does not seem at all relevant to ours. It is technically about decision-making, but not in the context of bias (other than present-bias). Curious for other opinions on this.

**5. Shivakumar (2018) (Cricket)**

Cricket introduced Umpiring decision review system (DRS): works in a similar way as the hawk-eye in tennis (correct challenge gets reinstated). One of the valuable things I found was that it seems that generated a lot of debate, criticism, skepticism among fans in newspapers and magazines,etc. Their results indicate an improvement in the fairness game though.

**6. Can Awareness Reduce (and Reverse) Identity-driven Bias in Judgement? Evidence from International Cricket\* - Chowdhury, Jewel, Singleton (2023)**

The paper focuses on how awareness on potential systematic biases from referees can help reduce and even reverse biases in judgement.

They are claiming that awareness and scrutiny of judgment bias can not only eliminate bias, but reverse the direction of it.

The setting they are using is very particular though, because in International cricket, they used to use umpires that shared the nationality of the home team. That changed when policy makers decided to employ neutral country empires. What the authors analyze then is the return of those umpires that shared the nationality during COVID times.

So, for umpires that used to benefit the home team in cricket, now they find that they actually benefit the visitor’s team. The rationale being overcompensating behavior. By being heavily scrutinized and making the previous biases well-known and highlighted in the media, it seems that umpires tend to benefit the visitor’s team in close judgments.

They use the machinery of Koszegi and Szeidl to give structure of why awareness and scrutiny MAY enable a reversal in the umpire’s preferences (decisions to where to give the call)

*One of the prescribed tools to deter judgement bias is to raise awareness, make its existence public, and encourage or enable more scrutiny of the officials.*

I think it’s a compelling idea to think that umpires that gained consciousness on their systematic biases now are making more equilibrated decisions, and that in case of erring, they are doing it in the opposite direction.

In relation to our paper, I think we could discuss how the VAR gives the referee a second chance to review their original decision and that in the process they may gain awareness of why they made the call. It could be for example that a particular play looked like a foul and that he recognizes that the audience’s pressure was the final trigger he needed to call it.

I think a good way to cite this paper in our intro would be by saying that Chowdhury et al stresses the importance of awareness of own biases when it comes to balancing the judgments. It doesn’t seem completely clear that awareness of referee home biases is what VAR brings to the table, but rather a second chance to review the decision and the fact that there are other referees scrutinizing the plays in a room where presumably the crowd noise plays a smaller part on their judgment calls.

**5. Shivakumar (2018) (Cricket) and 6. Chowdhury (2023)** were suggested by Referee 1 saying that:

*There is also a difference in so far that the calls in cricket are mostly not subjective with the technology, but in football they still can be. I am surprised the paper here doesn’t discuss this. There could still be home advantage in the VAR adjudication, or when the referee reviews footage on the pitch side monitor after being referred to it. Maybe it is worth discussing this*

**7. Ferraresi & Gucciardi (2022)**

They try to capture which effects dominates the performance of an individual in front of an audience: either the motivation (that enhances performance) or the fear of not meeting the expectations. They do it by analazing penalty kicks in the top 5 soccer leagues in just one season 2019/2020, exploiting the covid lockdown. They find that both a supportive audience and the size of it have a key positive role in performance:

Missing a penalty kick is more likely for home teams playing without their fans, while the contrary is true for away teams.

I’m guessing by the comments of the Referee that we could just squeeze the citation in the footnote 5 as suggested.

**8. Holder et al (2022):**

They only use 2 leagues and before-after comparison. Two seasons before (2015/2016 and 2016/2017) and two seasons after (2017/2018 and 2018/2019) the introduction of VAR.

Main conclusions are:

*In the case of the Italian Serie A, referees likely showed a preferential treatment of the home team, which manifested in added time decisions before the introduction of VAR, but this bias was not present in the two seasons after the introduction of the new system;*

*(iii) home teams were awarded more penalty kicks and fewer red cards, but, concerning penalties, VAR intervened equally often for both the home and away teams;*

Similar to our results.

*(iv) no differences were found when comparing the frequency of VAR interventions between experienced and inexperienced referees.*

Similar to our results.

“*it seems plausible that the home team gains an advantage from playing in front of the home crowd, but not as a result of referee favouritism”*

This is based on their results for red cards and penalty kicks, which are just looking at sample means for each year…

They mentioned in their conclusion:  
*Also, future studies could extend the considered time frame, include other leagues and record further game-specific characteristics.*

We find very similar results as them regarding , but with a more rigorous empirical strategy and considerably more data (leagues and years).

*Overall, these modest findings and even non-existent differences indicate that home bias occurs for reasons other than referees, suggesting that the process for training, promoting, and selecting referees at the highest league works well.*

Pretty big standard errors though, ours don’t show a difference in red cards and are way more precisely estimated.

**9. A systematic review of the literature on video assistant referees in soccer: Challenges and opportunities in sports analytics. Schmidt, de Oliveira, Steffen, Trojan (2023)**

In general, the authors cite that VAR hardly changes the games.

Not sure what to make of it.

Richard: I do not think this paper is relevant at all. It gives just one paragraph of “literature review” and then just talks about how much referees have to travel… It is also very poorly written and organized.

**10. Dufner et al (2023)**

The introduction of the Video Assistant Referee supports the fairness of the game – An analysis of the home advantage in the German Bundesliga

We probably want to cite this study, because is trying to tackle the same question with a higher ambition: fairness. Issues: One league, before and after comparison.

They claim in their abstract that referees are said to be the number 1 cause of Home Advantage, but they don’t provide any sort of proof or cites of that.

This is interesting, they even mentioned:  
*However, as Benz and Lopez (2021) point out, generalization based on findings from a single league may not be applicable. Future studies should thus examine the effects found in other soccer leagues.*

So we can write something like: Dufner et al (2023) and Benz and Lopez (2021) point out the necessity of a generalization of their findings, by a study that examines the effects on many leagues.

Among the strengths they cite in their paper, relative to other work such as Holder et al (2021) or Lago-Peñas et al (2021), is that it provides a *larger dataset over several seasons from a specific league.*  This is still only one league, so a more nuanced argument of the external validity is missing in the paper.

Richard: I agree with Camilo that we should cite them, even though it is a very bad study. I think their point about one-league studies relates more to the external validity rather than the inability to account for trends in home field advantage. The biggest weaknesses here are the fact that they only use one league, they include COVID-affected seasons in the treated group even though they acknowledge they are a bad comparison, and the empirical methodology is just really bad. Table 3 says that HFA goes away after the implementation of VAR because d > 0.2, which seems like an arbitrary cutoff, even though the coefficient only decreases by 18% (probably because of COVID-19) and is still VERY statistically significant (they treat their 1224 observations as IID, which is a bad choice). While some studies \*effectively\* compare averages, this one \*literally\* compares averages before and after implementation. The only thing this paper has going for it is that it does try to look directly at HFA in addition to fouls, cards, penalties, etc.

Richard again: I’m going to say this is a null effect of VAR on HFA. Also says decrease in fouls and no change in red cards/penalties.

They say the following, which gets at our point on how referee bias might not exist: “future studies should investigate whether some indicators of referee bias actually reflect correct decisions to different playing styles or changing levels of aggression” ie it is the players, not the refs, that cause more fouls and yellows to be called on the away team. Cites Buraimo (2010) and Hill & Van Yperen (2021), might be worth looking into those and potential citing them in our discussion/conclusion.

**11. Singleton et al (2023)**

**A decade of violence and empty stadiums in Egypt:**

Effects of crowds on behavior and decision making: Two questions being answered:

Does reinstating a supportive crowd after a long period of absence affect performances?

Is any reduced HFA from competing in empty stadiums robust to the repetition of this experiment? (This is because they banned crowds in Egypt, then they reinstated and then covid came so crows were removed again)

Among the conclussions: Results suggest that not all home teams are benefited from having their home crowds, perhaps because of the particular context of Egypt, where the authors suggest that the degree of social pressure was once extremely high but it may have cooled down because of the bans. (masculine fans groups and ultras movements have been dissolved). The channel is:

Less powerful crowds -> less pressure on the ref and other officials performance

Shivakumar (2015):

* Shows how implementation of a review system can impact the original decision-making of referees
* Cites literature on home umpire bias
* Other barely relevant; does not really discuss home field advantage at all

Chowdhury et al (2023):

* Cricket used to use home umpires (umpires from the country hosting a game). They were biased, so cricket started using neutral umpires (and home field advantage presumably decreased). During COVID-19, cricket went back to home umpires because of travel restrictions and such.
* Punchline: home field advantage decreased even more relative to the pre-pandemic period
* Basically, home umpires were aware of the scrutiny resulting from their previous bias and overcorrected, instead showing bias towards away teams instead of home teams